

# Convergence of Runge-Kutta Methods Applied to Linear Partial Differential-Algebraic Equations

K. Debrabant <sup>a,\*</sup> K. Strehmel <sup>b</sup>

<sup>a</sup>*Darmstadt University of Technology, Department of Mathematics,  
Schloßgartenstraße 7, D-64283 Darmstadt, Germany*

<sup>b</sup>*Martin-Luther-Universität Halle-Wittenberg, FB Mathematik and Informatik,  
Institut für Numerische Mathematik, Postfach, D-06099 Halle (Saale), Germany*

---

## Abstract

We apply Runge-Kutta methods to linear partial differential-algebraic equations of the form  $A u_t(t, x) + B(u_{xx}(t, x) + ru_x(t, x)) + Cu(t, x) = f(t, x)$ , where  $A, B, C \in \mathbb{R}^{n,n}$  and the matrix  $A$  is singular. We prove that under certain conditions the temporal convergence order of the fully discrete scheme depends on the time index of the partial differential-algebraic equation. In particular, fractional orders of convergence in time are encountered. Furthermore we show that the fully discrete scheme suffers an order reduction caused by the boundary conditions. Numerical examples confirm the theoretical results.

*Key words:* Partial differential-algebraic equations, Coupled systems, Implicit Runge-Kutta methods, Convergence estimates

---

## 1 Introduction

In this paper we consider linear partial differential-algebraic equations (PDAEs) of the form

$$A u_t(t, x) + B (u_{xx}(t, x) + ru_x(t, x)) + C u(t, x) = f(t, x), \quad (1)$$

---

\* Corresponding author.

*Email addresses:* `debrabant@mathematik.tu-darmstadt.de` (K. Debrabant),  
`strehmel@mathematik.uni-halle.de` (K. Strehmel).

where  $t \in (t_0, t_e)$ ,  $x \in \Omega = (-l, l) \subset \mathbb{R}$ ,  $A, B, C \in \mathbb{R}^{n,n}$  are constant matrices,  $r \in \mathbb{R}$ ,  $u, f : [t_0, t_e] \times \overline{\Omega} \rightarrow \mathbb{R}^n$ . We are interested in cases where the matrix  $A$  is singular. The singularity of  $A$  leads to the differential-algebraic aspect. It will always be tacitly assumed that the exact solution is as often differentiable as the numerical analysis requires.

In contrast to parabolic initial boundary value problems with regular matrices  $A$  and  $B$ , here we cannot prescribe initial and boundary values for all components of the solution vector, they have to fulfill certain consistency conditions. We consider one example:

**Example 1** *Superconducting coil (see Marszalek/Trzaska, Campbell/Marszalek [8,2]):*

$$\begin{pmatrix} 0 & 0 \\ -\frac{LC}{l^2} & \frac{L}{D} \end{pmatrix} u_{tt} - u_{xx} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} u = 0$$

with  $x \in (0, l)$ ,  $t > 0$ .  $u_1(t, x)$  denotes the voltage,  $u_2(t, x)$  denotes the divergence of the electric field strength within the coil.  $l$  is the length of the whole winding.  $L$ ,  $C$  and  $D$  are further coil parameters. Transformation to a partial differential-algebraic system of first order in  $t$  yields

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{LC}{l^2} & \frac{L}{D} \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} u_t - \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} u_{xx} + \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} u = 0. \quad (2)$$

As initial conditions we choose

$$u_1(0, x) = \left( \frac{E}{l} - \frac{CDEl}{6} \right) x + \frac{CDE}{6l} x^3, \quad u_3(0, x) = 0$$

and as boundary conditions

$$u_1(t, 0) = u_2(t, 0) = 0, \quad u_1(t, l) = E, \quad u_2(t, l) = CDE,$$

where  $E$  is the energizing source voltage at the input of the coil.

As the boundary values of  $u_1$  and  $u_2$  are constant, we get from the third and fourth equation of (2) that  $u_3$  and  $u_4$  fulfill homogeneous boundary conditions. From the initial condition of  $u_1$  and the first equation we derive  $u_2(0, x) = \frac{CDE}{l}x$ . With  $u_3(0, x) = 0$  and the third equation it follows  $u_{1t}(0, x) = 0$  and therefore  $u_{1xxt}(0, x) = 0$ . With the first equation this implies  $u_{2t}(0, x) = 0$ , and with the fourth equation we get finally  $u_4(0, x) = 0$ .

Here we have chosen the prescribed initial and boundary values such that all initial and boundary values are compatible.

For further examples considering the determination of the initial and boundary values which cannot be prescribed see Lucht/S./Eichler-Liebenow [7].

In the following we assume that for the numerical computation all initial values

$$u(t_0, x) = \varphi(x), \quad x \in \bar{\Omega},$$

and all boundary values entering into the space discretization are known,

$$Bu(t, x) = \psi(t, x), \quad x \in \partial\Omega, \quad t \in [t_0, t_e],$$

where we restrict ourselves to Dirichlet boundary conditions to simplify the presentation.

Investigations of the convergence of Runge-Kutta methods applied to abstract parabolic differential equations can be found for example in Brenner/Crouzeix/Thomée [1], Lubich/Ostermann [6] and Ostermann/Thalhammer [10]. The approach used there cannot be carried forward directly to the class of problems considered here because the matrix  $A$  is singular.

This paper is organized as follows: In Section 2 we derive a semi-discrete system based on finite differences. The result is a method-of-lines-DAE (MOL-DAE).

Section 3 is devoted to the Runge-Kutta approximation of the MOL-DAE. Under a regular transformation, the MOL-DAE of dimension  $nN$  is decoupled into  $N$  systems of dimension  $n$ , where  $N$  denotes the number of grid points on the  $x$ -axis. Furthermore, a Weierstrass-Kronecker transformation is used to decouple each of these systems into an ODE-system and an algebraic system. We introduce the differential time index of the linear PDAE and give the Runge-Kutta approximation to these subsystems.

In Section 4 we prove the convergence of  $L$ -stable Runge-Kutta discretizations with constant step sizes. The attained order of convergence in time depends on the differential time index of the PDAE and on the boundary conditions (homogeneous or inhomogeneous) which enter into the space discretization. Numerical experiments are finally presented in Section 5. We illustrate our convergence results for the backward Euler method and the 3-stage Radau IIA method.

## 2 Space discretization

The discretization in space of problem (1) by means of finite-differences results in a differential-algebraic equation (MOL-DAE)

$$M\dot{U} = DU(t) + \tilde{F}(t), \quad t_0 \leq t \leq t_e, \quad (3)$$

where  $U(t)$  is an  $Nn$ -dimensional real vector consisting of approximations to  $u$  at the grid points. Here  $N$  denotes the number of grid points on the  $x$ -axis. The matrix  $M$  is given by  $M = I_N \otimes A$  and the matrix  $D$  originates from the discretization of the differential operator  $B \frac{\partial^2}{\partial x^2}$  by second order difference-approximations, from the discretization of the differential operator  $rB \frac{\partial}{\partial x}$  by second ( $\delta = \frac{1}{2}$ ) or first order difference-approximations ( $\delta \in [0, 1] \setminus \{\frac{1}{2}\}$ ) and from the matrix  $C$ , i.e.,  $D$  is given by

$$D = -\frac{1}{h^2}P \otimes B - I_N \otimes C,$$

where  $I_N$  is the  $N$ -dimensional identity matrix,

$$P = \begin{pmatrix} -(2 - hr(1 - 2\delta)) & 1 + hr\delta & & \\ 1 + hr(\delta - 1) & -(2 - hr(1 - 2\delta)) & 1 + hr\delta & \\ \dots & \dots & \dots & \\ \dots & \dots & 1 + hr(\delta - 1) & -(2 - hr(1 - \delta)) \end{pmatrix}$$

and  $h = \frac{2l}{N+1}$  denotes the constant grid size. The  $Nn$ -dimensional real vector  $\tilde{F}(t)$  arises from the right hand side  $f$  of (1) and the boundary values which enter into the discretization.

We denote by  $U_h(t)$  the restriction of  $u(t, x)$  to the spatial grid and by  $\alpha_h(t)$  the space truncation error defined by

$$\alpha_h(t) := M\dot{U}_h(t) - DU_h(t) - \tilde{F}(t). \quad (4)$$

By Taylor expansion of the exact solution we get

$$\alpha_h(t) = h^{p_x}(I_N \otimes B)\gamma_h(t) \quad \text{with} \quad \|\gamma_h(t)\|_\infty \leq K, \quad (5)$$

where  $p_x \in \{1, 2\}$  is the order of approximation of the space discretization and  $K$  is a positive constant, i.e.,

$$\alpha_h(t) = \mathcal{O}(h^{p_x}) \quad \text{as} \quad h \rightarrow 0.$$

Furthermore, we can show that there exists a regular matrix  $Q$  with

$$Q^{-1} \frac{1}{h^2} P Q = \text{diag}\{\lambda_1, \dots, \lambda_N\}, \quad (6)$$

where

$$\lambda_j = -\frac{2 - hr(1 - 2\delta)}{h^2} + 2\frac{1 + hr\delta}{h^2} \sqrt{\frac{1 + hr(\delta - 1)}{1 + hr\delta}} \cos \frac{j\pi}{N+1}.$$

In the discrete  $L_2$ -norm we have

$$\max\{\|Q\|, \|Q^{-1}\|\} \leq C_1 \quad (7)$$

with a positive constant  $C_1$  independent of  $h$ . Therefore, in the following this norm is used.

### 3 Runge-Kutta approximations

In order to numerically advance in time the solution of the MOL-DAE (3), we employ an  $s$ -stage Runge-Kutta method

$$U_{m+1}^{(i)} = U_m + \tau \sum_{i=1}^s a_{ij} K_{m+1}^{(i)}, \quad MK_{m+1}^{(i)} = DU_{m+1}^{(i)} + \tilde{F}(t_m + c_i \tau), \quad 1 \leq i \leq s,$$

$$U_{m+1} = U_m + \tau \sum_{i=1}^s b_i K_{m+1}^{(i)},$$

where  $a_{ij}, b_i, c_i \in \mathbb{R}$  are the coefficients of the method and  $\tau = \frac{t_e - t_0}{M_e}$  the time step size.

For the investigation of the convergence of the method, it is useful to introduce the Runge-Kutta matrix  $\mathfrak{A} = (a_{ij})_{ij=1}^s$  and the vector notation  $\mathbb{1}_s = (1, \dots, 1)^\top \in \mathbb{R}^s$ ,  $b = (b_1, \dots, b_s)^\top$ .

Then, with the Kronecker product, we obtain the compact scheme

$$U_{m+1} = U_m + \tau (b^\top \otimes I_{Nn}) K_{m+1}, \quad (8a)$$

$$S_{m+1} = \mathbb{1}_s \otimes U_m + \tau (\mathfrak{A} \otimes I_{Nn}) K_{m+1}, \quad (8b)$$

$$(I_s \otimes M) K_{m+1} = (I_s \otimes D) S_{m+1} + \bar{F}(t_{m+1}), \quad m = 0, \dots, M_e - 1, \quad (8c)$$

where  $S_{m+1} = \left( U_{m+1}^{(1)\top}, \dots, U_{m+1}^{(s)\top} \right)^\top$ ,  $K_{m+1} = \left( K_{m+1}^{(1)\top}, \dots, K_{m+1}^{(s)\top} \right)^\top$  and  $\bar{F}(t_{m+1}) = \left( \tilde{F}(t_m + c_1 \tau)^\top, \dots, \tilde{F}(t_m + c_s \tau)^\top \right)^\top$ .

By the regular transformation (6), the MOL-DAE (3) can be decoupled into  $N$  DAEs

$$A \dot{U}_{Qk}(t) = D_k U_{Qk}(t) + \tilde{F}_{Qk}(t), \quad k = 1, \dots, N, \quad (9)$$

with  $D_k = -\lambda_k B - C$  and

$$\left( U_{Q1}(t)^\top, \dots, U_{QN}(t)^\top \right)^\top = (Q^{-1} \otimes I_n) U(t),$$

$$\left( \tilde{F}_{Q1}(t)^\top, \dots, \tilde{F}_{QN}(t)^\top \right)^\top = (Q^{-1} \otimes I_n) \tilde{F}(t).$$

In the following we assume that the matrix pencil  $\{D + \lambda M\}$ ,  $\lambda \in \mathbb{C}$ , is regular, which is equivalent to the regularity of all the matrix pencils  $\{D_k + \lambda A\}$ .

**Definition 2** Suppose that all matrix pencils  $\{D_k + \lambda A\}$ ,  $k = 1, \dots, N$ , are

regular and have the same index  $\nu_{dt}$ . Then the differential time index of the linear PDAE (1) is defined to be  $\nu_{dt}$ .

According to Weierstrass and Kronecker there exist regular matrices  $P_k$  and  $Q_k$  with

$$P_k A Q_k = \text{diag}\{I_{n_{k1}}, \dots, I_{n_{ks_k}}, N_{m_{k1}}, \dots, N_{m_{kl_k}}\}, \quad (10a)$$

$$P_k D_k Q_k = \text{diag}\{R_{k1}, \dots, R_{ks_k}, I_{m_{k1}}, \dots, I_{m_{kl_k}}\}, \quad (10b)$$

where

$$R_{ki} = \begin{pmatrix} \varkappa_{ki} & 1 & & 0 \\ & \ddots & \ddots & \\ & & \varkappa_{ki} & 1 \\ 0 & & & \varkappa_{ki} \end{pmatrix} \in \mathbb{C}^{n_{ki}, n_{ki}}, \quad N_{m_{ki}} = \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & 0 & 1 \\ 0 & & & 0 \end{pmatrix} \in \mathbb{C}^{m_{ki}, m_{ki}} \quad (10c)$$

(see Hairer/Wanner [5]), and for the differential time index of the PDAE it follows

$$\nu_{dt} = \max_k \{m_{ki} : i = 1, \dots, l_k\}.$$

Therefore, DAE (9) is decoupled into systems of the form

$$\dot{U}_{1kl}(t) = R_{kl} U_{1kl}(t) + \tilde{F}_{1kl}(t), \quad l = 1, \dots, s_k, \quad (11a)$$

$$N_{m_{kl}} \dot{U}_{2kl}(t) = U_{2kl}(t) + \tilde{F}_{2kl}(t), \quad l = 1, \dots, l_k \quad (11b)$$

with

$$\left( U_{1k1}(t)^\top, \dots, U_{1ks_k}(t)^\top, U_{2k1}(t)^\top, \dots, U_{2kl_k}(t)^\top \right)^\top = Q_k^{-1} U_{Qk}$$

and

$$\left( \tilde{F}_{1k1}(t)^\top, \dots, \tilde{F}_{1ks_k}(t)^\top, \tilde{F}_{2k1}(t)^\top, \dots, \tilde{F}_{2kl_k}(t)^\top \right)^\top = P_k \tilde{F}_{Qk}.$$

Similarly, DAE (4) can be transformed to

$$\dot{U}_{h1kl}(t) = R_{kl} U_{h1kl}(t) + \tilde{F}_{1kl}(t) + \alpha_{h1kl}(t), \quad l = 1, \dots, s_k, \quad (12a)$$

$$N_{m_{kl}} \dot{U}_{h2kl}(t) = U_{h2kl}(t) + \tilde{F}_{2kl}(t) + \alpha_{h2kl}(t), \quad l = 1, \dots, l_k. \quad (12b)$$

Runge-Kutta methods are invariant under the transformations (6) and (10). Therefore, to analyze convergence it is sufficient to apply them to systems of the form (11a) and (11b). Application to (11a) yields

$$U_{1kl,m+1} = U_{1kl,m} + \tau \left( b^\top \otimes I_{n_{kl}} \right) K_{1kl,m+1}, \quad (13a)$$

$$S_{1kl,m+1} = \mathbb{1}_s \otimes U_{1kl,m} + \tau (\mathfrak{A} \otimes I_{n_{kl}}) K_{1kl,m+1}, \quad (13b)$$

$$K_{1kl,m+1} = (I_s \otimes R_{kl}) S_{1kl,m+1} + \bar{F}_{1kl}(t_{m+1}), \quad m = 0, \dots, M_e - 1, \quad (13c)$$

and to (11b)

$$U_{2kl,m+1} = U_{2kl,m} + \tau \left( b^\top \otimes I_{n_{kl}} \right) K_{2kl,m+1}, \quad (14a)$$

$$S_{2kl,m+1} = \mathbb{1}_s \otimes U_{2kl,m} + \tau (\mathfrak{A} \otimes I_{n_{kl}}) K_{2kl,m+1}, \quad (14b)$$

$$(I_s \otimes N_{m_{kl}}) K_{2kl,m+1} = S_{2kl,m+1} + \bar{F}_{2kl}(t_{m+1}), \quad m = 0, \dots, M_e - 1. \quad (14c)$$

Now we start our convergence investigations.

#### 4 Convergence estimates

At first we introduce the global (space-time discretization) error  $e_{m+1}$  and the residual (space-time discretization) errors  $\delta_{m+1}$  and  $\Delta_{m+1}$  at the time level  $t = t_{m+1}$ .

**Definition 3** *The global error  $e_{m+1}$  at  $t_{m+1}$  is defined by*

$$e_{m+1} := U_h(t_{m+1}) - U_{m+1}$$

*and the residual errors  $\delta_{m+1}, \Delta_{m+1}$  are given by*

$$\delta_{m+1} := U_h(t_m + \tau) - U_h(t_m) - \tau \left( b^\top \otimes I_{Nn} \right) \hat{K}_{m+1}, \quad (15a)$$

$$\Delta_{m+1} := \hat{S}_{m+1} - \mathbb{1}_s \otimes U_h(t_m) - \tau (\mathfrak{A} \otimes I_{Nn}) \hat{K}_{m+1}, \quad (15b)$$

*where  $\hat{S}_{m+1}$  and  $\hat{K}_{m+1}$  are defined by the exact solution  $U_h(t)$  of the PDAE, i.e.,*

$$\begin{aligned} \hat{S}_{m+1} &:= \left( U_h(t_m + c_1\tau)^\top, \dots, U_h(t_m + c_s\tau)^\top \right)^\top, \\ \hat{K}_{m+1} &:= \left( \dot{U}_h(t_m + c_1\tau)^\top, \dots, \dot{U}_h(t_m + c_s\tau)^\top \right)^\top. \end{aligned}$$

**Definition 4** *The discretization scheme (8) is convergent of order  $(p_x, p^*)$ , if the global error satisfies*

$$\|e_{m+1}\| = \mathcal{O}(h^{p_x}) + \mathcal{O}(\tau^{p^*}) \quad \text{for} \quad (m+1)\tau = \text{const.}, \quad \tau, h \rightarrow 0,$$

*whenever  $u(t, x)$  is sufficiently often differentiable.*

With the components  $e_{Qk,m+1}$  defined by

$$\left( e_{Q1,m+1}^\top, \dots, e_{QN,m+1}^\top \right)^\top := (Q^{-1} \otimes I_n) e_{m+1}$$

and (7) we obtain the estimate

$$\frac{1}{C_1} \sqrt{h \sum_{k=1}^N \|e_{Qk,m+1}\|^2} \leq \|e_{m+1}\| \leq C_1 \sqrt{h \sum_{k=1}^N \|e_{Qk,m+1}\|^2}. \quad (16)$$

Letting  $\alpha_{h1kl,m+1} = \left( \alpha_{h1kl}(t_m + c_1\tau)^\top, \dots, \alpha_{h1kl}(t_m + c_s\tau)^\top \right)^\top$  we get with (12a) and (13c)

$$\hat{K}_{1kl,m+1} - K_{1kl,m+1} = (I_s \otimes R_{kl}) \left( \hat{S}_{1kl,m+1} - S_{1kl,m+1} \right) + \alpha_{h1kl,m+1}.$$

Using (13b), the transformed components  $e_{1kl,m} = U_{h1kl}(t_m) - U_{1kl,m}$  of the global discretization error and the transformed components  $\Delta_{1kl,m+1}$  of (15b) we obtain

$$\hat{S}_{1kl,m+1} - S_{1kl,m+1} = \mathbb{1}_s \otimes e_{1kl,m} + \tau (\mathfrak{A} \otimes I_{n_{kl}}) \left( \hat{K}_{1kl,m+1} - K_{1kl,m+1} \right) + \Delta_{1kl,m+1}.$$

Combining the last two equations leads to

$$G(\tau R_{kl}) \left( \hat{K}_{1kl,m+1} - K_{1kl,m+1} \right) = (\mathbb{1}_s \otimes R_{kl}) e_{1kl,m} + (I_s \otimes R_{kl}) \Delta_{1kl,m+1} + \alpha_{h1kl,m+1}, \quad (17)$$

where  $G(z) = I_s - z\mathfrak{A}$ .

**Remark 5** For a (matrix-valued) function  $f(z) : \mathbb{C} \rightarrow \mathbb{C}^{m,n}$  which is analytic in a neighbourhood of  $\kappa_{ki}$ , the matrix function  $f(R_{ki})$  with  $R_{ki}$  given in (10c) is defined by (see Golub/van Loan [4])

$$f(R_{ki}) := \begin{pmatrix} f_{i_1 i_2}(\mathfrak{x}_{ki}) \frac{f_{i_1 i_2}^{(1)}(\mathfrak{x}_{ki})}{1!} & \cdots & \frac{f_{i_1 i_2}^{(n_{ki}-1)}(\mathfrak{x}_{ki})}{(n_{ki}-1)!} \\ & \ddots & \vdots \\ & f_{i_1 i_2}(\mathfrak{x}_{ki}) & \frac{f_{i_1 i_2}^{(1)}(\mathfrak{x}_{ki})}{1!} \\ 0 & & f_{i_1 i_2}(\mathfrak{x}_{ki}) \end{pmatrix}_{i_1=1,\dots,m, \ i_2=1,\dots,n}.$$

In the following we assume that the Runge-Kutta method is A-stable and  $\Re(\mathfrak{x}_{kl}) \leq 0$  or  $|\kappa_{kl}| \leq C_2$  for all  $h \in (0, h_0]$  with a positive constant  $C_2$ . Then for sufficiently small  $\tau$  the matrix  $G(\tau R_{kl})$  is regular, and the Runge-Kutta system (13) has a unique solution. Using (13a), (17) and the transformed components  $\delta_{1kl,m+1}$  of (15a) we obtain the recursion

$$e_{1kl,m+1} = R(\tau R_{kl}) e_{1kl,m} + L(\tau R_{kl}) \Delta_{1kl,m+1} + \tau J(\tau R_{kl}) \alpha_{h1kl,m+1} + \delta_{1kl,m+1} \quad (18)$$

for the discretization error  $e_{1kl,m+1}$ , where we have used the abbreviations

$$J(z) = b^\top G(z)^{-1}, \quad R(z) = 1 + J(z) \mathbb{1}_s z, \quad L(z) = J(z) z$$

( $R(z)$  equals the classical stability function of the Runge-Kutta method).



Solving the recursion (18) with  $e_0 = 0$  leads to

$$e_{1kl,m+1} = \sum_{i=0}^m R(\tau R_{kl})^i L(\tau R_{kl}) \Delta_{1kl,m+1-i} + \tau \sum_{i=0}^m R(\tau R_{kl})^i J(\tau R_{kl}) \alpha_{h1kl,m+1-i} + \sum_{i=0}^m R(\tau R_{kl})^i \delta_{1kl,m+1-i}. \quad (19)$$

Now we assume that the Runge-Kutta method under consideration has (classical) order  $p$  and stage order  $q$  ( $p \geq q$ ). Then the simplifying conditions (see Hairer/Wanner [5])

$$B(p) : \sum_{i=1}^s b_i c_i^{k-1} = \frac{1}{k}, \quad k = 1, \dots, p,$$

$$C(q) : \sum_{j=1}^s a_{ij} c_j^{k-1} = \frac{1}{k} c_i^k, \quad i = 1, \dots, s, \quad k = 1, \dots, q,$$

are fulfilled.

With a Taylor expansion of  $U_h(t_m + c_j \tau)$  and  $\dot{U}_h(t_m + c_j \tau)$ ,  $j = 1, \dots, s$ , around  $t_m$  up to the order  $p$  we obtain for the  $j$ -th component of the residual error  $\Delta_{m+1}$  the equation

$$\Delta_{j,m+1} = \sum_{r=q+1}^p \frac{\tau^r}{r!} \left( \tilde{c}^r - r \mathfrak{A} \tilde{c}^{r-1} \right)_j U_h^{(r)}(t_m) + r_{\Delta_j, m+1}, \quad \|r_{\Delta_j, m+1}\| = \mathcal{O}(\tau^{p+1})$$

with  $\tilde{c}^i = (c_1^i, \dots, c_s^i)^\top$ . Therefore, with  $r_{\Delta 1kl, m+1} = \left( r_{\Delta 1kl, m+1}^\top, \dots, r_{\Delta s1kl, m+1}^\top \right)^\top$  and

$$W_r(z) = \frac{L(z) [\tilde{c}^r - r \mathfrak{A} \tilde{c}^{r-1}]}{1 - R(z)},$$

the error equation (19) can be written as

$$\begin{aligned} e_{1kl,m+1} &= \tau \sum_{i=0}^m R(\tau R_{kl})^i J(\tau R_{kl}) \alpha_{h1kl,m+1-i} + \sum_{i=0}^m R(\tau R_{kl})^i \delta_{1kl,m+1-i} \\ &\quad + \sum_{i=0}^m R(\tau R_{kl})^i L(\tau R_{kl}) r_{\Delta 1kl, m+1} \\ &\quad + \underbrace{\sum_{i=0}^m R(\tau R_{kl})^i (I_{n_{kl}} - R(\tau R_{kl})) \sum_{r=q+1}^p \frac{\tau^r}{r!} W_r(\tau R_{kl}) U_{h1kl}^{(r)}(t_{m-i})}_{=\kappa}. \end{aligned} \quad (20)$$

**Remark 6** The function  $W_r(z)$  was introduced by Ostermann/Roche [9] to investigate the convergence of Runge-Kutta methods for abstract scalar parabolic differential equations.

For the subsequent error estimate, the term  $\kappa$  is transformed in the following manner: By exchanging the order of summation we get

$$\kappa = \sum_{r=q+1}^p \frac{\tau^r}{r!} \left( \sum_{i=0}^{m-1} R(\tau R_{kl})^{m-i} W_r(\tau R_{kl}) U_{h1kl}^{(r)}(t_i) + W_r(\tau R_{kl}) U_{h1kl}^{(r)}(t_m) - \sum_{i=1}^m R(\tau R_{kl})^{m-i+1} W_r(\tau R_{kl}) U_{h1kl}^{(r)}(t_i) - R(\tau R_{kl})^{m+1} W_r(\tau R_{kl}) U_{h1kl}^{(r)}(t_0) \right).$$

From this we obtain

$$\kappa = \sum_{r=q+1}^p \frac{\tau^r}{r!} \left( \sum_{i=0}^{m-1} R(\tau R_{kl})^{m-i} W_r(\tau R_{kl}) \left( U_{h1kl}^{(r)}(t_i) - U_{h1kl}^{(r)}(t_{i+1}) \right) + W_r(\tau R_{kl}) U_{h1kl}^{(r)}(t_m) - R(\tau R_{kl})^{m+1} W_r(\tau R_{kl}) U_{h1kl}^{(r)}(t_0) \right).$$

Therefore it holds

$$\kappa = \sum_{r=q+1}^p \frac{\tau^r}{r!} \left( - \sum_{i=0}^{m-1} R(\tau R_{kl})^{m-i} W_r(\tau R_{kl}) \int_{t_i}^{t_{i+1}} U_{h1kl}^{(r+1)}(s) ds + W_r(\tau R_{kl}) U_{h1kl}^{(r)}(t_m) - R(\tau R_{kl})^{m+1} W_r(\tau R_{kl}) U_{h1kl}^{(r)}(t_0) \right).$$

A similar transformation can be found in Brenner/Crouzeix/Thomée [1].

Inserting this into (20) results in

$$\begin{aligned} e_{1kl,m+1} &= \tau \sum_{i=0}^m R(\tau R_{kl})^i J(\tau R_{kl}) \alpha_{h1kl,m+1-i} + \sum_{i=0}^m R(\tau R_{kl})^i \delta_{1kl,m+1-i} \\ &\quad + \sum_{i=0}^m R(\tau R_{kl})^i L(\tau R_{kl}) r_{\Delta 1kl,m+1} \\ &\quad + \sum_{r=q+1}^p \frac{\tau^r}{r!} \left( - \sum_{i=0}^{m-1} R(\tau R_{kl})^{m-i} W_r(\tau R_{kl}) \int_{t_i}^{t_{i+1}} U_{h1kl}^{(r+1)}(s) ds \right. \\ &\quad \left. + W_r(\tau R_{kl}) U_{h1kl}^{(r)}(t_m) - R(\tau R_{kl})^{m+1} W_r(\tau R_{kl}) U_{h1kl}^{(r)}(t_0) \right). \end{aligned}$$

Assuming that the Runge-Kutta matrix  $\mathfrak{A}$  is regular we can derive an analogous equation for the components  $e_{2kl,m+1}$  of the transformed global discretiza-

tion error

$$\begin{aligned}
e_{2kl,m+1} = & \tau \sum_{i=0}^m \tilde{R}(N_{kl})^i \tilde{J}(N_{kl}) \alpha_{h2kl,m+1-i} + \sum_{i=0}^m \tilde{R}(N_{kl})^i \delta_{2kl,m+1-i} \\
& + \sum_{i=0}^m \tilde{R}(N_{kl})^i \tilde{L}(N_{kl}) r_{\Delta 2kl,m+1} \\
& + \sum_{r=q+1}^p \frac{\tau^r}{r!} \left( - \sum_{i=0}^{m-1} \tilde{R}(N_{kl})^{m-i} \tilde{W}_r(N_{kl}) \int_{t_i}^{t_{i+1}} U_{h2kl}^{(r+1)}(s) ds \right. \\
& \left. + \tilde{W}_r(N_{kl}) U_{h2kl}^{(r)}(t_m) - \tilde{R}(N_{kl})^{m+1} \tilde{W}_r(N_{kl}) U_{h2kl}^{(r)}(t_0) \right)
\end{aligned}$$

with the abbreviations

$$\begin{aligned}
\tilde{J}(z) &= b^\top (I_s z - \tau \mathfrak{A})^{-1}, & \tilde{R}(z) &= 1 + \tau \tilde{J}(z) \mathbb{1}_s, \\
\tilde{L}(z) &= \tau \tilde{J}(z), & \tilde{W}_r(z) &= \frac{\tilde{L}(z) [\tilde{c}^r - r \mathfrak{A} \tilde{c}^{r-1}]}{1 - \tilde{R}(z)}.
\end{aligned}$$

Finally, using (5), we get for  $e_{Qk,m+1} = Q_k(e_{1k1,m+1}^\top, \dots, e_{1ks_k,m+1}^\top, e_{2k1,m+1}^\top, \dots, e_{2kl_k,m+1}^\top)^\top$  the equation

$$\begin{aligned}
e_{Qk,m+1} = & h^{p_x} \sum_{j=1}^s \tau \sum_{i=0}^m Q_k \text{diag}\{\dots, R(\tau R_{kj_1})^i J_j(\tau R_{kj_1}), \dots, \\
& \tilde{R}(N_{m_{kj_2}})^i \tilde{J}_j(N_{m_{kj_2}}), \dots\} P_k B \gamma_{hQk}(t_{m-i} + c_j \tau) \\
& + \sum_{i=0}^m Q_k \text{diag}\{\dots, R(\tau R_{kj_1})^i, \dots, \tilde{R}(N_{m_{kj_2}})^i, \dots\} Q_k^{-1} \delta_{Qk,m+1-i} \\
& + \sum_{j=1}^s \sum_{i=0}^m Q_k \text{diag}\{\dots, R(\tau R_{kj_1})^i L_j(\tau R_{kj_1}), \dots, \\
& \tilde{R}(N_{m_{kj_2}})^i \tilde{L}_j(N_{m_{kj_2}}), \dots\} Q_k^{-1} r_{\Delta_j Q, m+1} \\
& + \sum_{r=q+1}^p \frac{\tau^r}{r!} Q_k \left( \text{diag}_k\{\dots, W_r(\tau R_{kj_1}), \dots, \tilde{W}_r(N_{m_{kj_2}}), \dots\} Q_k^{-1} U_{Qk}^{(r)}(t_m) \right. \\
& - \sum_{i=0}^{m-1} \int_{t_i}^{t_{i+1}} \text{diag}_k\{\dots, R(\tau R_{kj_1})^{m-i} W_r(\tau R_{kj_1}), \dots, \tilde{R}(N_{m_{kj_2}})^{m-i} \tilde{W}_r(N_{m_{kj_2}}), \dots\} \\
& Q_k^{-1} U_{Qk}^{(r+1)}(s) ds \\
& \left. - \text{diag}_k\{\dots, R(\tau R_{kj_1})^{m+1} W_r(\tau R_{kj_1}), \dots, \tilde{R}(N_{m_{kj_2}})^{m+1} \tilde{W}_r(N_{m_{kj_2}}), \dots\} \right. \\
& \left. Q_k^{-1} U_{Qk}^{(r)}(t_0) \right). \tag{21}
\end{aligned}$$

Now we can estimate the different terms in (21). For that purpose we assume in the following that the matrix norms

$$\|Q_k \text{diag}\{N_{n_{k1}}^i, \mathbf{o}, \dots, \mathbf{o}\} Q_k^{-1}\|, \dots, \|Q_k \text{diag}\{\mathbf{o}, \dots, N_{m_{kl_k}}^i\} Q_k^{-1}\| \quad (22a)$$

and

$$\|Q_k \text{diag}\{N_{n_{k1}}^i, \mathbf{o}, \dots, \mathbf{o}\} P_k B\|, \dots, \|Q_k \text{diag}\{\mathbf{o}, \dots, \mathbf{o}, N_{m_{kl_k}}^i\} P_k B\| \quad (22b)$$

are bounded for  $i = 0, \dots, \max\{\nu_{dt}, n_{kj_1} : j_1 = 1, \dots, s_k\} - 1$  and all  $h \in (0, h_0]$ , where  $\mathbf{o}$  denotes a zero matrix.

Because of the  $A$ -stability of the Runge-Kutta method and  $\Re(\kappa_{kj_1}) \leq 0$  or  $|\kappa_{kj_1}| \leq C_2$  for all  $h \in (0, h_0]$  we have that  $\|R(\tau R_{kj_1})^i\|$ ,  $\|J_j(\tau R_{kj_1})\|$  and  $\|L_j(\tau R_{kj_1})\|$  are bounded for sufficiently small  $\tau$ . We assume further that  $R(it) \neq 1$  for  $t \in \mathbb{R} \setminus \{0\}$  and  $\lim_{z \rightarrow -\infty} R(z) \neq 1$ . Then  $W_r(\tau R_{kj_1})$  exists and is bounded. Moreover, as it is shown in Ostermann/Roche [9], one has

$$\|\tau^r W_r(\tau R_{kj_1})\| = \mathcal{O}(\tau^{\min\{p, q+2+\alpha\}}) \|R_{kj_1}^{\max\{0, \min\{p-r, q+2+\alpha-r\}\}}\|$$

with  $\alpha \in \mathbb{R}$ ,  $\alpha \geq -1$ .

Assuming that  $|\kappa_{kj_1}| \leq C_3(1 + |\lambda_k|)$  one can show (cf. D.[3], the proof relies on the Mean Value Theorem and Abel's partial summation formula) that

$$\|R_{kj_1}^{1+\alpha}\| \|U_{Q_k}^{(r)}(t_m)\| = k^{1+2\alpha} \mathcal{O}(h^{-\frac{1}{2}}).$$

Altogether the terms in (21) that originate from  $e_{1kl, m+1}$  are of order

$$\mathcal{O}(h^{p_x}) + \mathcal{O}(\tau^{\min\{p, q+2+\alpha\}}) h^{-\frac{1}{2}} k^{1+2\alpha}. \quad (23)$$

With the Taylor expansion

$$\gamma_{hQk}(t_{m-i} + c_j \tau) = \sum_{l=0}^{\nu_{dt}-2} \sum_{k=0}^l \frac{\tau^l (-i)^{l-k}}{k!(l-k)!} c_j^k \frac{\partial^l}{\partial t^l} \gamma_{hQk}(t_m) + \mathcal{O}(\tau^{\nu_{dt}-1}),$$

the term

$$a = h^{p_x} \sum_{j=1}^s \tau \sum_{i=0}^m Q_k \text{diag}\{0, \dots, 0, \tilde{R}(N_{m_{kj_2}})^i \tilde{J}_j(N_{m_{kj_2}}), 0, \dots, \} P_k B \gamma_{hQk}(t_{m-i} + c_j \tau)$$

of  $e_{Qk,m+1}$  can be written as

$$a = \sum_{j=0}^{\nu_{dt}-1} \frac{1}{j!} \sum_{l=0}^{\nu_{dt}-2} Q_k \text{diag}\{0, \dots, 0, \frac{N_{m_{kj_2}}^j}{\tau^{j-l}} \sum_{k=0}^l \frac{1}{k!(l-k)!} \sum_{i=0}^m (-i)^{l-k} (\tau^{j+1} \tilde{R}(z)^i \tilde{J}(z) \tilde{c}^k)^{(j)}(0), 0, \dots, 0\} P_k B \frac{\partial}{\partial t^l} \gamma_{hQk}(t_m) \\ + \sum_{j=0}^{\nu_{dt}-1} \frac{1}{j!} \sum_{i=0}^m Q_k \text{diag}\{0, \dots, 0, N_{m_{kj_2}}^j (\tau^{j+1} \tilde{R}(z)^i \tilde{J}(z))^{(j)}(0), 0, \dots, 0\} P_k B \mathcal{O}(\tau^{\nu_{dt}-1}),$$

where  $(\dots)^{(j)}$  denotes the  $j$ -th derivative w.r.t.  $z$ .

For  $L$ -stable Runge-Kutta methods with regular coefficient matrix  $\mathfrak{A}$  we have  $R(\infty) = 1 - b^\top \mathfrak{A}^{-1} \mathbb{1} = 0$  and therefore  $\tilde{R}(0) = 0$ . If the matrix norms in (22a) are bounded, then  $\|a\|$  is bounded if

$$\sum_{k=0}^l \frac{1}{k!(l-k)!} \sum_{i=0}^j (-i)^{l-k} (\tau^{j+1} \tilde{R}(z)^i \tilde{J}(z) \tilde{c}^k)^{(j)}(0) = 0 \quad (24)$$

for  $l = 0, \dots, j-1$ ,  $j = 1, \dots, \nu_{dt}-1$ .

The remaining terms in the equation (21) yield the classical order  $p_{\nu_{dt}}$  of the Runge-Kutta method applied to a linear DAE of index  $\nu_{dt}$  with constant coefficients. Thus, altogether we have

$$e_{Qk,m+1} = \mathcal{O}(h^{p_x}) + \mathcal{O}(\tau^{p_{\nu_{dt}}}) + \mathcal{O}(\tau^{\min\{p, q+2+\alpha\}}) h^{-\frac{1}{2}} k^{1+2\alpha}.$$

From (16) it follows that we have to choose  $\alpha$  such that  $\sum_{k=1}^N k^{2(1+2\alpha)}$  is bounded for  $N \rightarrow \infty$ . This implies  $\alpha = -\frac{3}{4} - \varepsilon$ ,  $\varepsilon > 0$ , and we have

$$\|e_{m+1}\| = \mathcal{O}(h^{p_x}) + \mathcal{O}(\tau^{\min\{p_{\nu_{dt}}, q+1.25-\varepsilon\}}).$$

If the derivatives of order  $(q+1)$  w.r.t. the time of the boundary conditions that enter into the space discretization are homogeneous, i.e.

$$B \frac{\partial^{q+1} u}{\partial t^{q+1}} = 0, \quad (25)$$

(12a) yields

$$R_{kl} U_{h1kl}^{(r)}(t) = U_{h1kl}^{(r+1)}(t) - \tilde{F}_{1kl}(t) + \alpha_{h1kl}(t),$$

and instead of (23) we obtain the order

$$\mathcal{O}(h^{p_x}) + \mathcal{O}(\tau^{\min\{p, q+2+\alpha\}}) h^{-\frac{1}{2}} k^{2\alpha-1},$$

which implies  $\alpha = \frac{1}{4} - \varepsilon$ , and therefore

$$\|e_{m+1}\| = \mathcal{O}(h^{p_x}) + \mathcal{O}(\tau^{\min\{p_{\nu_{dt}}, q+2.25-\varepsilon\}}).$$

Summarized, we have the following convergence result for smooth enough solutions  $u(t, x)$  of the PDAE ( $\max\{p + 1, p + \nu_{dt} - 1\}$  times differentiable with respect to  $t$  in  $[t_0, t_e]$  and  $p_x + 2$  times differentiable with respect to  $x$  in  $[-l, l]$ ):

**Theorem 7** *Let the following assumptions be fulfilled for  $h \rightarrow 0$  ( $N \rightarrow \infty$ ) and  $k = 1, \dots, N$ :*

- a) *for the matrix pencils  $D_k + \lambda A$  there exist Weierstrass-Kronecker decompositions according to (10), and the matrix norms in (22) are bounded,*
- b)  *$\Re(\kappa_{kj_1}) \leq 0$  or  $|\kappa_{kj_1}| \leq C_2$  for all  $h \in (0, h_0]$ ,*
- c)  *$|\kappa_{kj_1}| \leq C_3(1 + |\lambda_k|)$ ,*
- d) *if  $\nu_{dt} > 2$  then (24) is fulfilled for  $l = 0, \dots, j - 1$ ,  $j = 1, \dots, \nu_{dt} - 1$ .*

Furthermore let the Runge-Kutta method be of consistency order  $p$ , stage order  $q$  and  $L$ -stable (if  $\nu_{dt} = 0$  or  $\nu_{dt} = 1$ , it suffices  $A$ -stability with  $\lim_{z \rightarrow -\infty} R(z) < 1$ ) with a regular matrix  $\mathfrak{A}$  and  $R(it) \neq 1$  for  $t \in \mathbb{R} \setminus \{0\}$ . Let  $p_{\nu_{dt}}$  be the classical order of the Runge-Kutta method applied to a linear DAE of index  $\nu_{dt}$  with constant coefficients.

Then the discretization method (8) converges for linear PDAEs after  $\nu_{dt}$  time steps with the order  $(p_x, p^*)$  in the discrete  $L_2$ -norm in space and in the maximum norm in time with

$$p^* = \begin{cases} \min\{p_{\nu_{dt}}, q + 1.25 - \varepsilon\}: & \text{inhomog. boundary conditions according to (25)} \\ \min\{p_{\nu_{dt}}, q + 2.25 - \varepsilon\}: & \text{homog. boundary conditions according to (25)} \end{cases}$$

and  $\varepsilon > 0$  arbitrary small.

**Remark 8** (1) Stage order  $q \geq \nu_{dt} - 2$  implies condition d) in Theorem 7 for  $\nu_{dt} = 3$  or  $\nu_{dt} = 4$ .

(2) The assumptions on the Runge-Kutta method are fulfilled, e.g., for the Radau IIA and the Lobatto IIIC methods and in the case of  $\nu_{dt} \leq 1$  also for the implicit midpoint rule.

(3) If  $p \geq p^* + 1$ , then for  $L$ -stable Runge-Kutta methods with  $zL(1, z)$  bounded for  $\Re(z) \leq 0$ , the condition that (22a) is bounded can be replaced by the boundedness of the matrix norms

$$\left\| \frac{1}{\kappa_{k1}} Q_k \text{diag}\{N_{n_{k1}}^i, \mathbf{o}, \dots, \mathbf{o}\} Q_k^{-1} \right\|, \dots, \left\| \frac{1}{\kappa_{ks_k}} Q_k \text{diag}\{\mathbf{o}, \dots, \mathbf{o}, N_{n_{ks_k}}^i, \mathbf{o}, \dots, \mathbf{o}\} Q_k^{-1} \right\|, \quad (26a)$$

$$\left\| Q_k \text{diag}\{\mathbf{o}, \dots, \mathbf{o}, N_{m_{k1}}^i, \mathbf{o}, \dots, \mathbf{o}\} Q_k^{-1} \right\|, \left\| Q_k \text{diag}\{\mathbf{o}, \dots, N_{m_{kl_k}}^i\} Q_k^{-1} \right\|. \quad (26b)$$

(4) If we choose  $\varepsilon = 0$ , then we get for  $p^* < p_{\nu_{dt}}$

$$\|e_{m+1}\| = \mathcal{O}(h^{p^*}) + \mathcal{O}\left(\sqrt{|\ln h|}\tau^{p^*}\right).$$

**Remark 9** For a given Runge-Kutta method, Theorem 7 can be specialized. E.g., if we take the implicit Euler method, the resulting BTCS method is convergent of time order 1 for arbitrary time index, if only the conditions a) and b) of Theorem 7 are fulfilled. For the Radau IIA methods with  $s \geq 2$  stages we get

$$p^* = \begin{cases} \min\{s + 1.25 - \varepsilon, 2s - 1\}: & \nu_{dt} = 0, 1, \text{ inhomog. b.c.s according to (25)} \\ \min\{s + 2.25 - \varepsilon, 2s - 1\}: & \nu_{dt} = 0, 1, \text{ homog. b.c.s according to (25)} \\ s + 2 - \nu_{dt} & : \nu_{dt} \geq 2 \end{cases}$$

as temporal order of convergence, provided that the assumptions a)-d) of Theorem 7 are fulfilled.

## 5 Numerical examples

The numerical examples given below illustrate our convergence results. For the time integration we use the backward Euler method and the code RADAU5, which is a variable step size implementation of the 3-stage Radau IIA method, see Hairer/Wanner [5]. The Euler and Radau IIA methods are of great importance in applications.

**Example 10** The backward Euler method is given by the parameters

$$s = 1, \mathfrak{A} = (1), b = c = (1), p = q = 1.$$

We consider the linear PDAE

$$\underbrace{\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}}_{=A} u_t + \underbrace{\begin{pmatrix} 0 & 0 & -1 \\ 0 & -1 & -1 \\ 0 & 0 & 0 \end{pmatrix}}_{=B} u_{xx} + \underbrace{\begin{pmatrix} -1 & -1 & -1 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}}_{=C} u = f,$$

a coupled system of two parabolic equations and one algebraic equation, with  $x \in [-0.5, 0.5]$ ,  $t \in [0, 1]$ . The right-hand side, initial and Dirichlet boundary

values are chosen such that

$$u(t, x) = \begin{pmatrix} x(x-1) \sin(t) \\ x(x-1) \cos(t) \\ x(x-1)(e^t + t^5) \end{pmatrix}$$

is the exact solution. It holds

$$D_k = -\lambda_k B - C = \begin{pmatrix} 1 & 1 & 1 + \lambda_k \\ 0 & 1 + \lambda_k & \lambda_k \\ 0 & 0 & 1 \end{pmatrix}.$$

With

$$P_k = \begin{pmatrix} \lambda_k + 1 & -1 & -\lambda_k - \lambda_k^2 - 1 \\ 0 & 1 & -\lambda_k - 1 \\ 0 & 0 & 1 \end{pmatrix} \text{ and } Q_k = \begin{pmatrix} \frac{1}{\lambda_k + 1} & 0 & 0 \\ 0 & \frac{1}{\lambda_k + 1} & \frac{1}{\lambda_k + 1} \\ 0 & 0 & 1 \end{pmatrix},$$

we obtain the Weierstrass-Kronecker decomposition

$$P_k A Q_k = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad P_k D_k Q_k = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Therefore, the PDAE has differential time index 3, and the assumptions (22) are fulfilled. Remark 9 yields that the BTCS method is convergent after three steps of time order 1. This is confirmed by the numerical experiment, Table 1 shows the observed order of convergence in time at  $(x = 1, t_e = 1)$ . The nota-

$0.1\tau^{-1}$	$2^2$	$2^3$	$2^4$	$2^5$	$2^6$	$2^7$
$0.1h^{-1}$						
$2^2$	0.81	0.91	0.96	0.98	0.99	0.99
$2^3$	0.81	0.91	0.96	0.98	0.99	0.99
$2^4$	0.81	0.91	0.96	0.98	0.99	0.99

Table 1

Numerically observed order of convergence in the discrete  $L_2$ -norm.

tion of the first element 0.81 denotes the observed order when refining the grid from  $(h = 0.1/2^2, \tau = 0.1/2)$  to  $(h = 0.1/2^2, \tau = 0.1/2^2)$ , i.e.,  $0.81 = \log_2 \xi$ , where  $\xi$  denotes the ratio of the error with  $(h = 0.1/2^2, \tau = 0.1/2)$  to the



error with  $(h = 0.1/2^2, \tau = 0.1/2^2)$ . Furthermore, we see that a simultaneous refinement of  $h$  and  $\tau$  yields no order reduction.

**Example 11** We consider the 3-stage Radau IIA method with consistency order  $p = 5$  and stage order  $q = 3$ , and the linear PDAE

$$\underbrace{\begin{pmatrix} 0 & 2 & 0 \\ 1 & -1 & 0 \\ 1 & -1 & 0 \end{pmatrix}}_{=A} u_t + \underbrace{\begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}}_{=B} u_{xx} + \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{=C} u = f(t, x)$$

with  $x \in [-1, 1]$  and  $t \in [0, 1]$ . This example shows the dependence of the time order on the boundary values.

1. We choose the right-hand side such that

$$u(t, x) = \left( x^2 e^{-t}, x^2 e^{-\frac{1}{2}t}, x^2 \sin t \right)^\top$$

is the exact solution. Then we have inhomogeneous boundary values

$$u(t, \mp 1) = \left( e^{-t}, e^{-\frac{1}{2}t}, \sin t \right)^\top.$$

Furthermore it holds

$$D_k = \begin{pmatrix} \lambda_k & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \lambda_k - 1 \end{pmatrix}, \quad P_k A Q_k = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad P_k D_k Q_k = \begin{pmatrix} \frac{2\lambda_k}{-\lambda_k - \eta_k} & 0 & 0 \\ 0 & \frac{2\lambda_k}{-\lambda_k + \eta_k} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

with

$$P_k = \begin{pmatrix} \frac{\lambda_k - \eta_k}{4\lambda_k} & 1 & 0 \\ \frac{\lambda_k + \eta_k}{4\lambda_k} & 1 & 0 \\ 0 & \frac{1}{1 - \lambda_k} & \frac{1}{\lambda_k - 1} \end{pmatrix}, \quad Q_k = \begin{pmatrix} \frac{4\lambda_k}{(\lambda_k + \eta_k)\eta_k} & -\frac{4\lambda_k}{(\lambda_k - \eta_k)\eta_k} & 0 \\ -\frac{\lambda_k}{\eta_k} & \frac{\lambda_k}{\eta_k} & 0 \\ -\frac{\lambda_k}{\eta_k(\lambda_k - 1)} & \frac{\lambda_k}{\eta_k(\lambda_k - 1)} & 1 \end{pmatrix}, \quad \eta_k = \sqrt{\lambda_k^2 + 8\lambda_k}$$

$(|\lambda_k + 8| > \frac{1}{2})$  for  $N > 3$ , i.e.  $h < \frac{1}{2}$ .

The PDAE has therefore differential time index 1, and the conditions a)-c) of Theorem 7 are fulfilled which yields convergence of time order  $4.25 - \varepsilon$ . This is confirmed by the numerical experiment, see Table 2.

2. If instead the right-hand side is chosen such that

$$u(t, x) = \left( (x^2 - 1)e^{-t} + t^3 \cos^2 x, x^2 e^{-\frac{1}{2}t}, (x^2 - 1) \sin t + t^3 \sin^2 x \right)^\top$$

$0.1\tau^{-1}$	$2^1$	$2^2$	$2^3$
$0.2h^{-1}$			
$2^3$	4.27	4.28	4.30
$2^4$	4.26	4.26	4.26
$2^5$	4.26	4.26	4.25
$2^6$	4.26	4.26	4.25

Table 2

Numerically observed order of convergence in the discrete  $L_2$ -norm for inhomogeneous boundary values.

is the exact solution, then we have inhomogeneous boundary values

$$u(t, \mp 1) = \left( t^3 \cos^2 1, e^{-\frac{1}{2}t}, t^3 \sin^2 1 \right)^\top,$$

but the derivatives of  $Bu(t, x)$  of order 4 w.r.t. the time vanish. Therefore, we obtain the convergence order 5 in time, see Table 3.

$0.1\tau^{-1}$	$2^2$	$2^3$
$0.2h^{-1}$		
$2^3$	5.00	5.00
$2^4$	5.00	5.00
$2^5$	5.00	5.00
$2^6$	5.00	5.00

Table 3

Numerically observed order of convergence in the discrete  $L_2$ -norm for inhomogeneous boundary values where the derivatives of  $Bu$  of order 4 w.r.t. the time vanish.

**Example 12** We consider the 3-stage Radau IIA method and the linear PDAE (2) describing the superconducting coil. It holds

$$D_k = \begin{pmatrix} \lambda_k & -1 & 0 & 0 \\ 0 & \lambda_k & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad P_k A Q_k = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad P_k D_k Q_k = \begin{pmatrix} -\frac{i\lambda_k}{\sqrt{1-\lambda_k}} & 0 & 0 & 0 \\ 0 & \frac{i\lambda_k}{\sqrt{1-\lambda_k}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

with

$$P_k = \begin{pmatrix} -\frac{i}{\sqrt{1-\lambda_k}} - \frac{i\sqrt{1-\lambda_k}}{\lambda_k} & 1 & -1 \\ \frac{i}{\sqrt{1-\lambda_k}} & \frac{i\sqrt{1-\lambda_k}}{\lambda_k} & 1 & -1 \\ 0 & 0 & 1 & -\frac{1}{\lambda_k} \\ \frac{1}{\lambda_k} & 0 & 0 & 0 \end{pmatrix}, \quad Q_k = \begin{pmatrix} \frac{1}{2(1-\lambda_k)} & \frac{1}{2(1-\lambda_k)} & 0 & -\frac{\lambda_k}{1-\lambda_k} \\ \frac{\lambda_k}{2(1-\lambda_k)} & \frac{\lambda_k}{2(1-\lambda_k)} & 0 & -\frac{\lambda_k}{1-\lambda_k} \\ -\frac{i\lambda_k}{2(1-\lambda_k)^{\frac{3}{2}}} & \frac{i\lambda_k}{2(1-\lambda_k)^{\frac{3}{2}}} & -\frac{\lambda_k}{1-\lambda_k} & 0 \\ -\frac{i\lambda_k^2}{2(1-\lambda_k)^{\frac{3}{2}}} & \frac{i\lambda_k^2}{2(1-\lambda_k)^{\frac{3}{2}}} & -\frac{\lambda_k}{1-\lambda_k} & 0 \end{pmatrix}.$$

The coil PDAE has therefore differential time index 2, and the conditions of Theorem 7 (with the matrix norms (22a) replaced by (26)) are fulfilled which yields an order of convergence in time of 3.

This is confirmed by the numerical experiment, see Table 4.

$0.1\tau^{-1}$	$2^4$	$2^5$	$2^6$
$0.2h^{-1}$			
$2^2$	3.00	3.00	3.00
$2^3$	3.00	3.00	3.00
$2^4$	3.00	3.00	3.00

Table 4

Numerically observed order of convergence in the discrete  $L_2$ -norm for the coil PDAE.

## 6 Conclusion

The attention has here been restricted to a class of linear partial differential-algebraic equations. We have given convergence results in dependence on the type of boundary values and the time index. When the error is measured in the discrete  $L_2$ -norm over the whole domain, the convergence order in time of the Runge-Kutta method for a smooth solution is in general non-integer and smaller than the order expected for differential-algebraic equations of the same index. Some numerical examples were presented and confirm the theoretical convergence results.

The extension of the analysis to the case of space  $d$  dimensional linear partial differential-algebraic equations of the form

$$A u_t(t, \vec{x}) + \sum_{i=1}^d B_i (u_{x_i x_i}(t, \vec{x}) + r_i u_{x_i}(t, \vec{x})) + C u(t, \vec{x}) = f(t, \vec{x})$$

with  $\vec{x} = (x_1, \dots, x_d)^\top$  and a cuboid as domain is possible, see D. [3], but becomes rather technical and offers no new insight. Furthermore, the con-

sideration of periodic boundary values is also possible. Here we could show as temporal convergence order the order of an ordinary differential-algebraic equation. In the case of Neumann boundary conditions, the temporal convergence order lies in between the order obtained for Dirichlet- and the order obtained for periodic boundary conditions.

Future work in this area will be concerned with convergence investigations for semi-linear partial differential-algebraic equations.

## Acknowledgements

The authors are very grateful to the referee for his comments and fruitful suggestions.

## References

- [1] P. Brenner, M. Crouzeix, V. Thomée: Single step methods for inhomogeneous linear differential equations in Banach space. R.A.I.R.O. Anal. Numér. 16 (1982) 5-26.
- [2] S. L. Campbell, W. Marszalek: The Index of an Infinite Dimensional Implicit System. Mathematical and Computer Modelling of Dynamical Systems 5 (1999) 18-42.
- [3] K. Debrabant: Numerische Behandlung linearer und semilinearer partieller differentiell-algebraischer Systeme mit Runge-Kutta-Verfahren. Dissertation, Martin-Luther-Universität Halle-Wittenberg, 2004.
- [4] G. H. Golub, C. F. van Loan: Matrix Computations. Third Edition. The John Hopkins University Press. Baltimore, London, 1996.
- [5] E. Hairer, G. Wanner: Solving Ordinary Differential Equations II. Stiff and Differential - Algebraic Problems. Springer-Verlag Berlin. Heidelberg, 1996.
- [6] Ch. Lubich, A. Ostermann: Runge-Kutta Methods for parabolic equations and convolution quadrature. Math. Comp. 60 (1993) 105-131.
- [7] W. Lucht, K. Strehmel and C. Eichler-Liebenow: Indexes and special discretization methods for linear partial differential algebraic equations. BIT 39 (1999) 484-512.
- [8] W. Marszalek, Z. W. Trzaska: Analysis of implicit hyperbolic multivariable systems. Appl. Math. Modelling 19 (1995) 400-410.
- [9] A. Ostermann, M. Roche: Runge-Kutta Methods for Partial Differential Equations and Fractional Order of Convergence. Math. Comp. 59 (1992) 403-420.
- [10] A. Ostermann, M. Thalhammer: Convergence of Runge-Kutta methods for nonlinear parabolic equations, Appl. Numer. Math. 42 (2002) 367-380.